The HJ-Biplot Visualization of the Singular Spectrum Analysis Method

Alberto Oliveira da Silva (✉) [0000-0002-3496-6802] , Adelaide Freitas[0000-0002-4685-1615]

Department of Mathematics - University of Aveiro
Aveiro, 3810-193, Portugal
albertos@ua.pt

Abstract. Time series data usually emerge in many scientific domains. The extraction of essential characteristics of this type of data is crucial to characterize the time series and produce, for example, forecasts. In this work, we take advantage of the trajectory matrix constructed in the Singular Spectrum Analysis, as well as of its decomposition through the Principal Component Analysis via Partial Least Squares, to implement a graphical display employing the Biplot method. In these graphs, one can visualize and identify patterns in time series from the simultaneous representation of both rows and columns of such decomposed matrices. The interpretation of various features of the proposed biplot is discussed from a real-world data set.

Keywords: Singular Spectrum Analysis, NIPALS algorithm, Biplots.

1 Overview

Singular Spectrum Analysis (SSA) is a non-parametric method and a suitable tool to perform exploratory analysis on time series [6]. The Basic SSA schema is the version that deals with the description and identification of the structure of a one-dimensional real-valued time series. Basic SSA can be described as two successive stages: *decomposition* and *reconstruction*. The first one is subdivided into step 1, the *embedding*, and step 2, the *Singular Value Decomposition* (SVD), while the second consists of two other phases, the *grouping* and the *diagonal averaging*. The primary purpose is to decompose the original time series into the sum of a few interpretable components, such as trend, oscillatory shape (e.g., seasonality) which should be separated from a noise component [5].

For any matrix, the factorization given by SVD allows practical graphical representations of both rows and columns of the matrix employing biplots methods [2, 3]. Biplots provide easier interpretations, are much more informative than the traditional scatterplots, and might facilitate the work in the grouping step in SSA. Several types of biplots can be constructed depending on how the three factors identified by SVD are aggregated to obtain only two factors. Herein, the option is the biplot method proposed by Galindo [3], called HJ-biplot, which yields a simultaneous representation of both rows and columns of a matrix of interest with maximum quality [3].

The main objective of this paper is to propose a new exploratory procedure to visualize and identify patterns in the time series through the construction of an HJ-biplot from the results of the SVD step on the Basic SSA. Moreover, this work suggests an alternative approach to obtain the factorization referred in step 2 (first stage) based on the *Nonlinear Iterative Partial Least Squares* (NIPALS) algorithm [11] instead of the usual SVD method. Although it provides equivalent results concerning the singular vectors and the singular values, it empowers the SSA to deal with missing values in the data, without employing any imputation method, since NIPALS is a suitable tool to treat this problem [10, 12]. That occurs because, in each iteration of the NIPALS algorithm, only present data are considered in the regressions performed, ignoring the missing elements. This is equivalent to defining all missing points in the least squares objective function as zero.

The paper is organized as follows. In Section 2, we provide a short description of the theoretical background related to the methods involved in this work. In Section 3, we propose a biplot approach to the SSA method and some possible interpretations of it. In Section 4, we perform an application of the proposed technique by using real-world data set. Final conclusions are contained in Section 5.

2 Methods

2.1 Basic Singular Spectrum Analysis

The Basic SSA is a model-free tool used to recognize and identify the structure of a time series [5]. As before mentioned, it is composed of two complementary stages, as follows.

First Stage: Decomposition.
Consider a real-valued time series $\mathrm{Y} = (y_1, \dots, y_N)$ of length N. Let the integer value L $(1 < L < N)$ be the so-called *window length*, as well as $K = N - L + 1$. Hereupon, the *embedding* procedure, that is the first step of the Basic SSA, consists in representing Y in K lagged vectors, $\mathbf{x}_1, \dots, \mathbf{x}_K$, each one of size L (L-lagged vectors), i.e., $\mathbf{x}_j = (y_j, \dots, y_{j+L-1})$, $1 \leq j \leq K$. This sequence of K vectors forms the trajectory matrix $\mathbf{X} = [\mathbf{x}_1 : \dots : \mathbf{x}_K]$, that has as its columns the L-lagged vectors. Step 2, the SVD step, results in the singular value decomposition of the trajectory matrix. Consider that rank($\mathbf{X}$) is equal to d, and the matrix $\mathbf{S}$ is defined as the product $\mathbf{X}'\mathbf{X}$. So, the SVD of $\mathbf{X}$ is the decomposition in the form

$$\mathbf{X} = \sum_{i=1}^{d} \sqrt{\lambda_i}\mathbf{u}_i\mathbf{v}_i', \tag{1}$$

where λ_i, $i = 1, \dots, d$, are the eigenvalues of the matrix $\mathbf{S}$ arranged in decreasing order of magnitudes ($\lambda_i > 0$), $\{\mathbf{v}_1, \dots, \mathbf{v}_d\}$ is the orthonormal system of the eigenvectors of $\mathbf{S}$ associated with the eigenvalues $\lambda_1, \dots, \lambda_d$, and

$$\mathbf{u}_i = \mathbf{X}\mathbf{v}_\mathbf{i}/\sqrt{\lambda_i}. \tag{2}$$

The elements of the triple $\sqrt{\lambda_i}, \mathbf{u}_i, \mathbf{v}_i$ are also known as *singular values*, *left* and *right singular vectors* of **X**, respectively. Besides, defining

$$\mathbf{X}_\mathrm{i} = \sqrt{\lambda_i}\mathbf{u}_i\mathbf{v}_i', \tag{3}$$

one can represent **X** as a sum of d 1-rank matrices, i.e.,

$$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_d. \tag{4}$$

Second Stage: Reconstruction.
Once the expansion (4) has been determined, the third step of the SSA starts with the partitioning of the index set $\{1, \dots, d\}$ into disjoints subsets $I_j, j = 1, \dots, p$. Let

$$\mathbf{X}_I = \sum_{i \in I} \mathbf{X}_i \tag{5}$$

and the decomposition can be written as

$$\mathbf{X} = \mathbf{X}_{I_1} + \cdots + \mathbf{X}_{I_p}. \tag{6}$$

The intention of the grouping procedure is the separation of the additive components of the time series [6]. The objective of the next phase, the *diagonal averaging* step, is to transform each matrix of the *grouping* decomposition into a new time series [5]. At this point, as in [6], it is convenient to define: $\mathbb{M}_{L,K}$ as the space of the matrices of dimension $(L \times K)$; $\mathbb{M}_{L,K}^{(H)}$ the space of Hankel matrices of dimension $(L \times K)$; the embedding operator $\mathcal{T}: \mathbb{R}^N \mapsto \mathbb{M}_{L,K}$ as $\mathcal{T}(Y) = \mathbf{X}$; and the projector $\mathcal{H}$ of $\mathbb{M}_{L,K}$ to $\mathbb{M}_{L,K}^{(H)}$, that carries out the projection by changing entries on auxiliary diagonals (where $i + j$ is a constant) to their averages along the diagonal. So, the diagonal averaging procedure corresponds to obtaining

$$\tilde{Y}^{(k)} = \mathcal{T}^{-1}[\mathcal{H}(\mathbf{X}_{I_k})] \tag{7}$$

and, then

$$Y = \sum_{k=1}^{p} \tilde{Y}^{(k)}. \tag{8}$$

2.2 PCA through NIPALS

The NIPALS algorithm belongs to the Partial Least Squares family, a set of iterative algorithms that implement a wide range of multivariate explanatory and exploratory techniques. The NIPALS is designed as an iterative estimation method for Principal Component Analysis (PCA), that computes the principal components through an iterative sequence of simple ordinary least squares regressions [10, 11]. It produces a singular value decomposition (SVD) of a matrix regardless of its dimensions and the presence of missing data [10]. Again, considering that the trajectory matrix has rank d, the method decomposes **X** as a sum of d 1-rank matrices in terms of the outer product of two vectors, a score $\mathbf{t}_i$ and a loading $\mathbf{p}_i$, so that

$$\mathbf{X} = \mathbf{t}_1\mathbf{p}_1' + \cdots + \mathbf{t}_d\mathbf{p}_d'. \quad (9)$$

The elements of the scores vector $\mathbf{t}_i$ are the projections of the sample points on the principal component direction, while each loading in $\mathbf{p}_i$ is the cosine of the angle between the component direction vector and a variable axis [4]. The NIPALS first computes $\mathbf{t}_1$ and $\mathbf{p}_1$ from $\mathbf{X}$ and, then, the outer product $\mathbf{t}_1\mathbf{p}_1'$ is subtracted from $\mathbf{X}$ to calculate the residual matrix $\mathbf{E}_1$. After, $\mathbf{E}_1$ is used to compute $\mathbf{t}_2$ and $\mathbf{p}_2$, and the residual $\mathbf{E}_2$ is calculated subtracting $\mathbf{t}_2\mathbf{p}_2'$ from $\mathbf{E}_1$, and so on until to obtain $\mathbf{t}_d$ and $\mathbf{p}_d$. The NIPALS algorithm is shown in Algorithm 1.

Algorithm 1. NIPALS internal relations.

NIPALS

Input: $\mathbf{E_0} = \mathbf{X}$
Output: $\mathbf{P} = |\mathbf{p_1}: \ldots: \mathbf{p}_d|, \mathbf{T} = |\mathbf{t_1}: \ldots: \mathbf{t}_d|$
for all $\boldsymbol{i} = \mathbf{1}, \ldots, \boldsymbol{d}$ **do**
step 0: initialize $\mathbf{t}_i$
step 1:
repeat
step 1.1: $\mathbf{p}_i = \mathbf{E}_{i-1}'\mathbf{t}_i/\mathbf{t}_i'\mathbf{t}_i$
step 1.2: $\mathbf{p}_i = \mathbf{p}_i/\|\mathbf{p}_i\|$
step 1.3: $\mathbf{t}_i = \mathbf{E}_{i-1}\mathbf{p}_i$
until convergence of $\mathbf{p}_i$
step 2: $\mathbf{E}_i = \mathbf{E}_{i-1} - \mathbf{t}_i\mathbf{p}_i'$
end for

From the internal relations in each iteration of the NIPALS algorithm, and after normalizing $\mathbf{t}_i$, such that

$$\mathbf{t}_i^* = \mathbf{t}_i/\|\mathbf{t}_i\| \Leftrightarrow \mathbf{t}_i = \sqrt{\mathbf{t}_i'\mathbf{t}_i}\,\mathbf{t}_i^*, \quad (10)$$

the following equations can be verified [10]:

$$\mathbf{E}_{i-1}'\mathbf{E}_{i-1}\mathbf{p}_i = \lambda_i\mathbf{p}_i \quad (11)$$

$$\mathbf{E}_{i-1}\mathbf{E}_{i-1}'\mathbf{t}_i^* = \lambda_i\mathbf{t}_i^*, \quad (12)$$

where $\lambda_i = \mathbf{t}_i'\mathbf{t}_i$ is the eigenvalue of both matrices $\mathbf{E}_{i-1}'\mathbf{E}_{i-1}$ and $\mathbf{E}_{i-1}\mathbf{E}_{i-1}'$, as well as $\mathbf{p}_i$ and $\mathbf{t}_i^*$ are their corresponding eigenvectors. Thus, the NIPALS decomposition of $\mathbf{X}$ can be written as

$$\mathbf{X} = \sqrt{\mathbf{t}_1'\mathbf{t}_1}\,\mathbf{t}_1^*\mathbf{p}_1' + \cdots + \sqrt{\mathbf{t}_d'\mathbf{t}_d}\,\mathbf{t}_d^*\mathbf{p}_d'. \quad (13)$$

Now, define the matrix $\mathbf{\Sigma}$ as a diagonal matrix containing the singular values $\sqrt{\mathbf{t}_i'\mathbf{t}_i}$ arranged in decreasing order. So, one can write the matrix form of the expansion (13) as

$$\mathbf{X} = \mathbf{T}^*\mathbf{\Sigma}\mathbf{P}', \quad (14)$$

where $\mathbf{T}^*$ is the scores matrix whose column vectors $\mathbf{t}_i^*$ are orthonormal, and $\mathbf{P}$ is the loadings matrix whose column vectors $\mathbf{p}_i$ are also orthonormal.

2.3 HJ-Biplot

The term biplot is due to Gabriel [2] and is associated to a graphical representation that reveals essential characteristics of multivariate data structure, e.g., patterns of correlations between variables or similarities between observations [7]. Consider a target data matrix $\mathbf{Z}$ of dimension $(I \times J)$, and its decomposition in the form

$$\mathbf{Z} = \mathbf{A}\mathbf{B}', \tag{15}$$

where $\mathbf{A}$ is a matrix of dimension $(I \times Q)$, and $\mathbf{B}$ is a matrix of dimension $(J \times Q)$. The matrices $\mathbf{A}$ and $\mathbf{B}$ create two sets of points, and if $Q = 2$, then the rows and columns of $\mathbf{Z}$ can be simultaneously represented into a two-dimensional graph called biplot, in which the rows of $\mathbf{A}$ are reproduced by points and the columns of $\mathbf{B}'$ are expressed as vectors connected to the origin (arrows). Thus, the biplot displays the row markers $\mathbf{a}_1, \dots, \mathbf{a}_I$ of $\mathbf{Z}$, as well as its column markers $\mathbf{b}_1, \dots, \mathbf{b}_J$, so that the inner product $\mathbf{a}_i'\mathbf{b}_j$ is the element z_{ij} of $\mathbf{Z}$ [8]. Very briefly, the interpretation of the biplot representation can be performed as follows:

1. The distance between points corresponds to how different the associated individuals are (dissimilarities), mainly if they are well represented;
2. The size of the arrow is proportional to the standard deviation of the associated variable. The longer the arrow, the greater the standard deviation;
3. The cosine of the angle between arrows approximates the correlation between the variables they represent. Thus, if the angle is next to 90° it indicates a poor correlation, while an angle close to 0° or 180° suggests a strong correlation, being positive in the first case and negative in the other.

The most popular biplot is the classic one [2], in which the metric of the columns is preserved. This version is also designated by GH-biplot [8]. An essential property of the GH-biplot is that the biplot vectors have the same configuration of the data matrix columns and the quality of representation of columns is maximum. By choosing row and column markers properly, the HJ-biplot allows representing the rows and columns simultaneously in the same Euclidean space with optimal quality for both [3].

To construct an HJ-biplot version based on NIPALS instead of SVD as proposed in [3], it's enough to demonstrate the relationship between $\mathbf{t}_i^*$ and $\mathbf{p}_i$, as will be done next.

From the equation (12), multiplying it to the left by $\mathbf{E}_{i-1}'$, it becomes

$$\mathbf{E}_{i-1}'\mathbf{E}_{i-1}(\mathbf{E}_{i-1}'\mathbf{t}_i^*) = \lambda_i(\mathbf{E}_{i-1}'\mathbf{t}_i^*) \tag{16}$$

Next, the vector normalization of $(\mathbf{E}_{i-1}'\mathbf{t}_i^*)$ results in $\mathbf{E}_{i-1}'\mathbf{t}_i^*/\sqrt{\mathbf{t}_i'\mathbf{t}_i}$, i.e., the vector $\mathbf{p}_i$. Proceeding in the same way with respect to equation (11), and multiplying it to the left by $\mathbf{E}_{i-1}$ we have

$$\mathbf{E}_{i-1}\mathbf{E}_{i-1}'(\mathbf{E}_{i-1}\mathbf{p}_i) = \lambda_i(\mathbf{E}_{i-1}\mathbf{p}_i). \tag{17}$$

After, $(\mathbf{E}_{i-1}\mathbf{p}_i)$ is normalized, which produces $\mathbf{E}_{i-1}\mathbf{p}_i/\sqrt{\mathbf{t}_i'\mathbf{t}_i}$, i.e., the vector $\mathbf{t}_i^*$. Hence,

$$\sqrt{\mathbf{t}_i'\mathbf{t}_i}\mathbf{p}_i = \mathbf{E}_{i-1}'\mathbf{t}_i^*, \tag{18}$$

and

$$\sqrt{\mathbf{t}_i'\mathbf{t}_i}\mathbf{t}_i^* = \mathbf{E}_{i-1}\mathbf{p}_i. \tag{19}$$

To unify the biplot axes scales similarly to what is done in [3], the following designation is done

$$\mathbf{a}_i = \mathbf{E}_{i-1}\mathbf{p}_i = \sqrt{\mathbf{t}_i'\mathbf{t}_i}\mathbf{t}_i^* \tag{20}$$

$$\mathbf{b}_i = \mathbf{E}_{i-1}'\mathbf{t}_i^* = \sqrt{\mathbf{t}_i'\mathbf{t}_i}\mathbf{p}_i. \tag{21}$$

Substituting (18) into (20), it follows that

$$\mathbf{a}_i = \mathbf{E}_{i-1}\mathbf{b}_i/\sqrt{\mathbf{t}_i'\mathbf{t}_i}, \tag{22}$$

and plugging (20) in (21) we get

$$\mathbf{b}_i = \mathbf{E}_{i-1}'\,\mathbf{a}_i/\sqrt{\mathbf{t}_i'\mathbf{t}_i}. \tag{23}$$

Thus, from (22) and (23), the coordinates of the *i*-th column are expressed as a function of the coordinates of the *i*-th row and vice versa. As a consequence, it allows the representation of the rows and columns in the same Cartesian coordinates system. Moreover, these expressions of the column and row coordinates lead to the maximum quality of the representation for rows and columns in the same system [3]. Considering the matrix form of the NIPALS decomposition in (14), it is worth to mention that for the configuration of the HJ-biplot, we have

$$\mathbf{A} = \mathbf{T}^*\mathbf{\Sigma}, \tag{24}$$

$$\mathbf{B} = \mathbf{P}\mathbf{\Sigma}, \tag{25}$$

and so,

$$\mathbf{X} \neq \mathbf{A}\mathbf{B}'. \tag{26}$$

3 The SSA-HJ-Biplot

The trajectory matrix that will be decomposed by the NIPALS algorithm at the second step of the first stage of the SSA has some peculiarities in relation to the usual multivariate data matrix. Instead of individuals and variables, the rows and columns of the trajectory matrix represent L-lagged and K-lagged vectors of a time series, respectively. That said, after the decomposition of $\mathbf{X}$, a row marker in the HJ-biplot denotes a K-lagged vector and is depicted in the graph as a point. In turn, a column marker repre-

sents a L-lagged vector, being that an arrow symbolizes it. One of the goals of SSA-HJ-biplot is to assist in the grouping step and for this, building more than one SSA-HJ-biplot may be needed. The first SSA-HJ-biplot uses the 1st and the 2nd principal components (PC), the next one uses the 2nd PC and the 3rd PC, and so on as long as the remain components can explain the variability of the data, which is given by

$$PC_{\%}^{(i)} = \mathbf{t}_i'\mathbf{t}_i / \sum_{j=1}^{d} \mathbf{t}_j'\mathbf{t}_j, \tag{27}$$

or visually through the scree plot of the singular values ($\sqrt{\mathbf{t}_i'\mathbf{t}_i}$) [5].

The window length L has to be large enough so that each L-lagged vector captures a substantial part of the behavior of the time series [5], but at the same time, it permits the interpretability of the graphics display. A window length equals to $N/2$ provides both capabilities because it allows for the most detailed decomposition [5]. The interpretation of the first SSA-HJ-biplot is performed in terms of:

1. The proximity of points. Biplot points whose Euclidean distances are small imply similarity in the behavior of the associated K-lagged vectors;
2. The length of the biplot vectors. If the arrows are roughly the same size, this indicates that the L-lagged vectors have standard deviation also close, which suggests that the process is stationary in the variance;
3. The angle formed between biplot vectors. If the angle between the two arrows is next to 0°, it hints a strong and positive autocorrelation between the two L-lagged vectors associated (negative if next to 180°). If the angle is close to 90°, it is expectable an autocorrelation near to zero.

It is worth to take in mind the percentage of explained variability represented by the first two components, since the higher the percentage, the better the quality of the adjust of the SSA-HJ-biplot [3].

As a rule, a singular value represents the contribution of the corresponding PC in the form of the time series. As the tendency generally characterizes the shape of a time series, its singular values are higher than the others, that is, they are the first eigenvalues [1]. On the other hand, when two singular values are close enough, i.e.,

$$\sqrt{\mathbf{t}_i'\mathbf{t}_i} \approx \sqrt{\mathbf{t}_h'\mathbf{t}_h},$$

this is an evidence of the formation of plateaus in the scree plot and indicates that the associated SSA-HJ-biplot is informative about the oscillatory components of the time series [5], as long as the PC explain high variability of the data.

4 Example

In this Section, an SSA-HJ-biplot is constructed to a time series that contains the records the carbon dioxide concentration in the Earth's atmosphere, measured monthly from January of 1965 to December of 1980 at an observing station on Mauna Loa in Hawaii [9], referred as T.S. CO2 in this work and that is represented in **Fig. 1**. Two auxiliary plots in **Fig. 2** provide some hints for what to expect in an SSA-HJ-biplot analysis in the data. In **Fig. 2** (b), where the 1st PC is plotted against an index $j =$

$1, \ldots, K$, the presence of a trend component in T.S. CO2 is manifest, and this should emerge somehow in the first SSA-HJ-biplot, i.e., in the biplot where the axes are the 1st and 2nd PCs. **Fig. 3** brings the first SSA-HJ-biplot, where one can verify that the 1st PC explains 67% of the data variability, i.e., the trend direction. A channel formed by two dotted lines helps in the perception of the presence of the trend, although the 2nd PC contributes to attenuate the slope if compared with the plot in **Fig. 2** (b). Each one of the biplot points (in red) represents a K-lagged vector, and its corresponding label indicates the month in which the lagged vector starts. In this sense, accordingly to the graph legend, a point labeled as "O" means a K-lagged vector starting in October of some year, and a label "D" symbolizes that the respective K-lagged vectors begins in December, and so on. These points are the row markers, determined by the rows of $\mathbf{T}^*\mathbf{\Sigma}$, that is $\mathbf{a}_i' = \mathbf{t}_i'$, $i = 1, \ldots, L$.

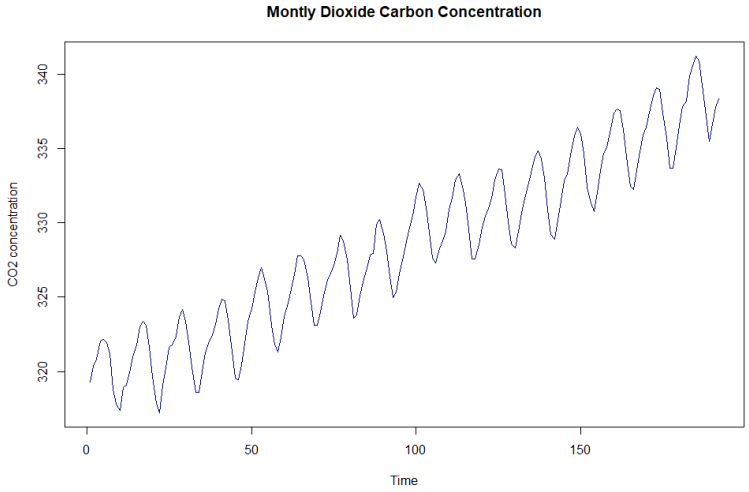

Fig. 1. Carbon dioxide concentration in the Earth's atmosphere measured monthly from January of 1965 to December of 1980 at an observing station on Mauna Loa in Hawaii.

According to the biplot theory, near points indicate similarity in the behavior of the lagged vectors, e.g., the points tagged as A, Y, and U in **Fig. 3**, i.e., the K-lagged vectors starting in April, May, and June. But not only that. Considering the labeling procedure before mentioned, the SSA-HJ-biplot is also capable of capturing the behavior of the months, since April, May, and June correspond precisely to the periods in which the highest concentration of carbon dioxide occurs in the atmosphere. It means that the points in the first SSA-HJ-biplot can represent not only the K-lagged vectors that start in a given month but also the month itself.

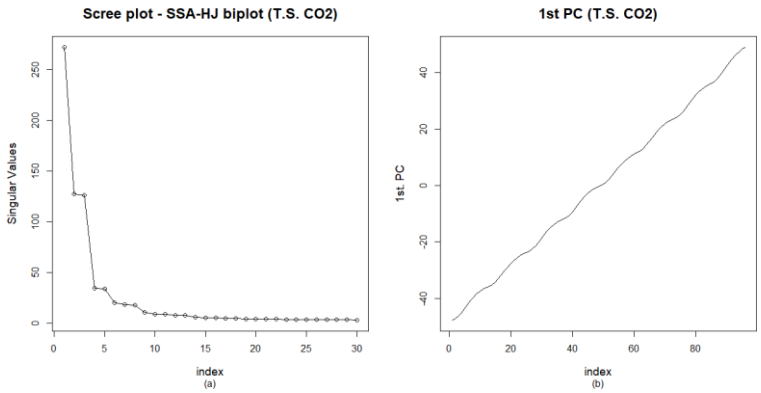

Fig. 2. Auxiliary plots in the SSA-HJ-biplot analysis.

In **Fig. 3**, the SSA-HJ-biplot represents the column markers (the L-lagged vectors) as black arrows up to the sixth L-lagged vector (tagged as $L1$ until $L6$), ordered from top to bottom. From the seventh L-lagged vector onwards the pattern repeats itself, and so they were plotted in gray. It means that the first group of arrows, which is at the top, refer to the L-lagged vectors beginning in January and July, just below those as starting in February and August, and so on. The angle between two consecutive arrows Li and Lj, such that $i = 1, \dots, 5$ and $j = i + 1$, indicates a strong autocorrelation between the respective L-lagged vectors since Li and Lj form very sharp angles. As for $L1$ and the others up to $L6$, the angles range from something close to 0 to something close to 90 degrees, which suggests a fading of the autocorrelations. And this cycle repeats from $L7$ periodically, which suggests the non-stationarity also in the seasonality.

Fig. 4 shows the SSA-HJ-biplot formed by the 2nd and 3rd PCs, while **Fig. 5** exhibit the SSA-HJ-biplot constructed from the 4th and 5th PCs. Along with the first SSA-HJ-biplot, these are the only ones that produce interpretable results or evidence some pattern in the time series, being that these results are in agreement with the one verified in the scree plot of the singular values in **Fig. 2** (a), where the pair of points related to $\sqrt{\mathbf{t}_2'\mathbf{t}_2}$ and $\sqrt{\mathbf{t}_3'\mathbf{t}_3}$ are around at the same level, the same with respect to $\sqrt{\mathbf{t}_4'\mathbf{t}_4}$ and $\sqrt{\mathbf{t}_5'\mathbf{t}_5}$. In the SSA-HJ-biplot of **Fig. 4**, there are well defined 12 groups of row markers, where each one of these groups refers to a K-lagged vector that starts for a specific month. Also, the column markers associated with each one of these groups show strong autocorrelation between the L-lagged vectors. All of this indicates a seasonal pattern, with peaks and valleys separated by 12 months. In turn, the SSA-HJ-biplot of **Fig. 5** groups the lagged vectors two by two, e.g., January and July, February and August, and so on. Interpreting this together with the biplot of **Fig. 4**, where these same groups occur but in the opposite directions, one can conclude that the valleys tend to be six months behind the peaks.

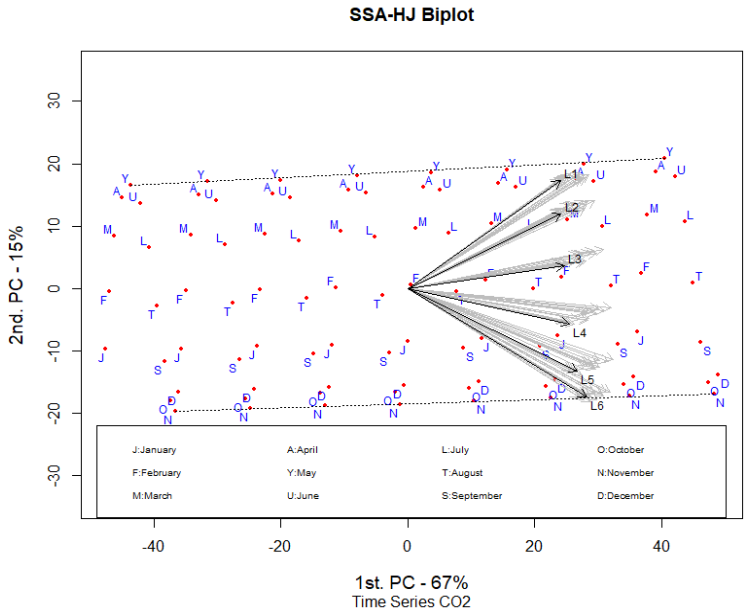

Fig. 3. First SSA-HJ-biplot of the T.S.CO2 trajectory matrix decomposition.

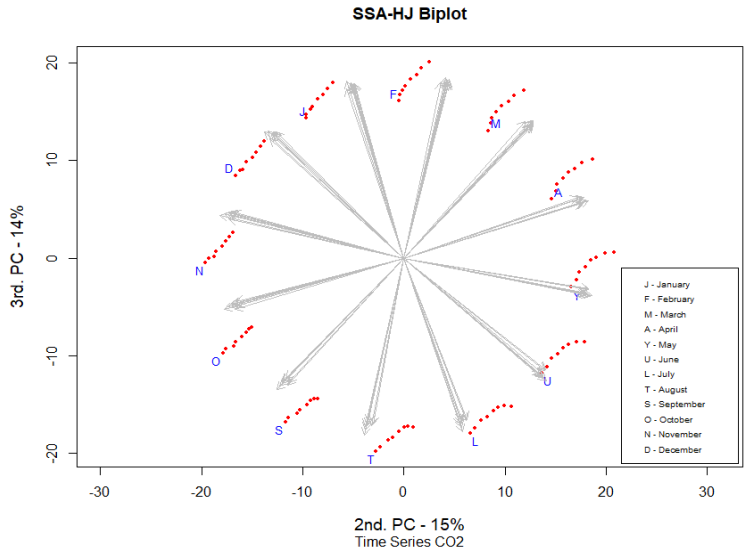

Fig. 4. The second SSA-HJ-biplot whose axes are the 2nd and 3rd PCs.

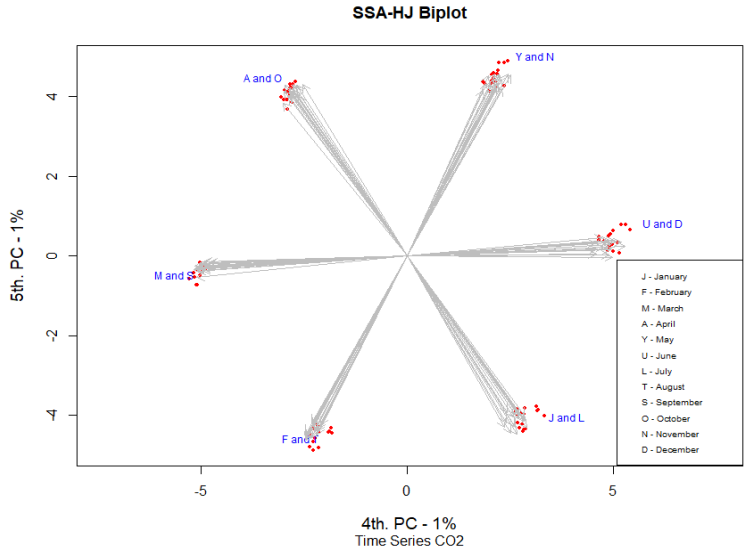

Fig. 5. The third SSA-HJ-biplot whose axes are the 4th and 5th PCs.

Therefore, the result of the grouping step for the decomposition of the T.S. CO2 should be $\mathbf{X}_1$ and $\mathbf{X}_2$, the first corresponding to the trend component, and the second describing the seasonal component, in which

$$\mathbf{X}_1 = \sqrt{\mathbf{t}_1'\mathbf{t}_1}\,\mathbf{t}_1^*\mathbf{p}_1', \tag{28}$$

and

$$\mathbf{X}_2 = \sum_{i=2}^{5}\sqrt{\mathbf{t}_i'\mathbf{t}_i}\,\mathbf{t}_i^*\mathbf{p}_i', \tag{29}$$

with the rest being related to the noise component.

5 Conclusions

This paper attempts to provide an alternative way to visualize and understand the underlying structure of the trajectory matrix, that is the result of the embedding step of the SSA. The HJ biplot visualization method appears to be a promisor exploratory technique adequate to the purposes of this work since it provides interpretability to the results of the SVD step as was illustrated by an application. The SSA-HJ-biplots and auxiliary graphics provided a visual solution for the decomposition of the analyzed time series, properly separating the trend and the oscillatory component, using biplot axes up to the fifth PC. Also, allowed the identification of all relevant eigentriple, composed by the singular values $\sqrt{\mathbf{t}_i'\mathbf{t}_i}$, by the left singular vectors $\mathbf{t}_i^*$, and by the right singular vectors $\mathbf{p}_i$, $i = 1, \ldots, 5$, to perform the grouping step. The study also revealed that the SSA-HJ-biplot points, representative of the row markers ($\mathbf{a}_i'$) and symbol of

the K-lagged vectors that begin in a given period of the series (months in this specific case) could also depict the period itself in terms of dissimilarities, being possible to visually verify the months with the highest and lowest levels of CO_2 concentration in the atmosphere throughout the years. The SSA-HJ-biplot built with the 1st and 2nd PCs proved yet to be useful in dealing with autocorrelations between the column markers, which are drawn as arrows and represent the L-lagged vectors. This study is promising in the sense that the SSA-HJ-biplot has a great potential as an exploratory tool to analyze the structure of a univariate time series due to its visual appeal in such a complex issue.

Acknowledgments.
The authors were supported by Fundação para a Ciência e a Tecnologia (FCT), within project UID/MAT/04106/2019 (CIDMA).

References

1. Alexandrov, T.: A method of trend extraction using Singular Spectrum Analysis. REVSTAT, Statistical Journal, **7**(1), 1-22 (2009)
2. Gabriel, K.: The biplot graphic display of matrices with application to principal component analysis. Biometrika, **58**(3), 453-467 (1971)
3. Galindo, M.P.: An alternative of simultaneous representation: HJ-biplot. Questiió, **10**(1), 13-23 (1986)
4. Geladi, P., Kowalsky, B.R.: Partial Least Squares regression: a tutorial. Analytica Chimica Acta, **185**, 1–17 (1986)
5. Golyandina, N., Nekrutkin, V., Zhigljavsky, A.: Analysis of Time Series Structure: SSA and Related Techniques. 1st ed. Chapman & Hall/CRC, Boca Raton, Florida (2001)
6. Golyandina, N., Shlemov, A.: Variations of Singular Spectrum Analysis for separability improvement: non-orthogonal decompositions of time series. Statistics and its Interface, **8**(3), 277–294 (2015)
7. Greenacre, M.: Biplots in Practice. FBBVA, Bilbao, Biscay (2010)
8. Nieto, A.B., Galindo, M.P., Leiva, V., Galindo, P.V.: A methodology for biplots based on bootstrapping with R. Colombian Journal of Statistics, **37**(2), 367–397 (2014)
9. NOAA Homepage, https://www.esrl.noaa.gov/gmd/ccgg/trends/, last accessed 2019/05/24
10. Vinzi, V.E., Russolillo, G.: Partial Least Squares algorithms and methods. WIREs Comput Stat, **5**, 1–19 (2013)
11. Wold, H.: Estimation of principal components and related models by iterative least squares. In: Krishnaiah, P.R. (ed.), Multivariate Analysis, New York: Academic Press, 391–420 (1966)
12. Wold, S., Albano, C., Dunn, W.J. III, Esbensen, K., Hellberg, S., Johansson, E., Sjostrom, M.: Pattern recognition: finding and using regularities in multivariate data. In: Martens, H., Russwurm, H. (eds.), Food Research and Data Analysis, London: Applied Science Publishers, 147–189 (1983)